GISLab Metadata Standard

FGDC Required Metadata to Accompany Spatial Data in GISLab Spatial Data Warehouse

Executive Order 12906 (Clinton, 1994) established the Federal Geographic Data Committee (FGDC), tasked with developing federal geospatial data standards. **Because all government agencies are required to adopt the FGDC standard, all data entering the GISLab spatial data warehouse must be accompanied by these metadata.** The outline below describes the content of the FGDC-required metadata fields. Please note that this outline highlights the *required* metadata fields; there are many more optional (but in some cases extremely useful) fields. The field descriptions and examples were excerpted from "Content Standard for Digital Geospatial Metadata Workbook (For use with FGDC-STD-001-1998), Version 2.0, Federal Geographic Data Committee, May 1, 2000." Refer to this document for more information on metadata structure and fields; a pdf version is available at http://www.gislab.lanl.gov/guides.html.

An empty version of this outline, an XML template, and examples of completed metadata for raster and vector datasets is available from the GISLab website (http://www.gislab.lanl.gov/policies.html). Simply fill in metadata under each **bold** field and mail or email the completed form (one per data set) to the Data Administrator, GISLab, EES-10, MS D452, LANL (gislab_data@lanl.gov).

Required Metadata Fields

The fields listed in **bold** below must be completed for each spatial data set. Some fields are automatically filled by ArcCatalog during data entry and are so noted on this outline.

1. Identification

General

Abstract

A brief narrative summary of the data set. The "Abstract" briefly describes the "what" aspects of the data set. For example, what information is in the data set? What area is covered?

Purpose

A summary of the intentions with which the data set was developed. The "Purpose" describes the "why" aspects of the data set. For example, why was the data set created?

Access Constraints

Restrictions and legal prerequisites for accessing the data set. These include any access constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on obtaining the data set.

Use Constraints

Restrictions and legal prerequisites for using the data set after access is granted. These include any use constraints applied to assure the protection of privacy or intellectual property, and any special restrictions or limitations on using the data set.

Contact

Contact information for an individual or organization that is knowledgeable about the data set.

Note that the “contact” meant here is the data owner/originator/steward, as distinct from the “metadata contact,” listed later, who prepared the metadata.

Details

Person

Organization

General

Contact Voice Telephone

Address

Address Type

E.g., "mailing," "physical," "mailing and physical," etc.

Address

City

State or Province

Postal Code

Citation

Details

General

Title (auto filled – file name)

The name by which the data set is known.

Originator

The name of an organization or individual that developed the data set. If the name of editors or compilers are provided, the name must be followed by "(ed.)" or "(comp.)" respectively. (The "Point of Contact" is the person(s) and organization(s) that can be contacted if questions arise about the data set. The "Originator" and the "Point of Contact" may be the same.)

Publication Date (YYYYMMDD)

The date when the data set is published or otherwise made available for release.

Time Period

Currentness Reference

The basis on which the time period of content information is determined. Information about the currentness of a data set (that is, information about how “up-to-date” is a data set) is important to many, if not most, potential users. Most users are interested in the currentness of a data set related to the “ground condition” (that is, when the “real world” looked the way it s described in the data set). Unfortunately, sometimes only the time that the information was recorded or published is known. The Currentness Reference element requires the producer to identify if the Time Period of Content dates and times refer to the ground condition, or some later time when the information was recorded, published, etc.

Calendar Date (YYYYMMDD)

The year (and optionally month, or month and day).

Status

Progress

The state of the data set; e.g., "Complete," "In work," or "Planned."

Update Frequency

The frequency with which changes and additions are made to the data set after the initial data set is completed; e.g., "Continually," "Daily," "Weekly," "Monthly," "Annually," “Unknown," "As needed," "Irregular," or "None planned."

Spatial Domain

Bounding Coordinates

The limits of coverage of a data set expressed by latitude and longitude values in the order western-most, eastern-most, northern-most, and southern-most. Geographic coordinates will be defined automatically using the projected coordinate datum (if previously defined for the dataset.) For data sets that include a complete band of latitude around the earth, the West Bounding Coordinate shall be assigned the value -180.0, and the East Bounding Coordinate shall be assigned the value 180.0.

North (auto filled if spatial reference has been applied to data)

South (auto filled if spatial reference has been applied to data)

East (auto filled if spatial reference has been applied to data)

West (auto filled if spatial reference has been applied to data)

Keywords

A keyword is a work or phrase that signifies the meaning or main ideas of a data set. They often are used as an index to the contents of a data set. The standards provide for four types of keywords: theme (the subject of the data set, such as wetlands, vegetation, etc.), place (the geographic location of the data set, such as Montgomery County, Yellowstone National Park), stratum (the vertical location of the data set, such as seafloor, seabed, troposphere, stratosphere), and temporal (time references for a data set, such as pre- Columbian, World War II).

Theme Keyword

Common-use word or phrase used to describe the subject of the data set.

Theme Thesaurus

Reference to a formally registered thesaurus or a similar authoritative source of theme keywords. For example, "Webster's Geographical Dictionary (1988)," "USGS, GNIS Digital Gazetteer (1994)— http://geonames.usgs.gov/," "NASA Thesaurus," or None.

Security Information

Handling restrictions imposed on the data set because of national security, privacy, or other concerns. (Use LANL security or classification descriptors.)

Security Classification System

Name of the classification system, e.g., "DOE"

Security Classification

Name of the handling restrictions on the data set, e.g., "Top secret," "Secret," "Confidential," "Restricted," "Unclassified," "Sensitive," etc.

Security Handling Description

Additional information about the restrictions on handling the data set.

2. Data Quality

General

Logical Consistency Report

A report on logical consistency shall describe the fidelity of relationships encoded in the data structure (topology) of the digital spatial data. For example, "Overshoots and undershoots exist in line intersections…," "data has/has not been 'cleaned/built'," or "Polygons intersecting the neatline are closed along the border. Segments making up the outer and inner boundaries of a polygon tie end-to-end to completely enclose the area. Line segments are a set of sequentially numbered coordinate pairs. No duplicate features exist nor duplicate points in a data string. Intersecting lines are separated into individual line segments at the point of intersection. Point data are represented by two sets of coordinate pairs, each with the same coordinate values…."

Completeness Report

Information about omissions, selection criteria, generalization, definitions used, and other rules used to derive the data set. For example, "All photo-interpretable wetlands are mapped. In the treeless prairies, 1/4 acre wetlands are mapped. In forested areas, small open water and emergent wetlands are mapped. In general, the minimum mapping unit is from 1 to 3 acres depending on the wetland type and the scale and emulsion of the source aerial photography…." (See example metadata at http://www.gislab.lanl.gov/policies/gislab_metadata.html.)

Attribute Accuracy

Attribute Accuracy Report

An explanation of the accuracy of the identification of the entities and assignments of values in the data set and a description of the tests used. For example, "Attribute accuracy is tested by manual comparison of the source with hard copy plots and/or symbolized display of the digital line graph on an interactive computer graphic system; selected attributes that cannot be visually verified on plots or on screen, are interactively queried and verified on screen…. Accuracy is estimated to be 98.5 percent," or "unknown," or "to be determined."

OR

Quantitative Attribute Accuracy Assessment

A value assigned to summarize the accuracy of the identification of the entities and assignments of values in the data set and the identification of the test that yielded the value.

Attribute Accuracy Value

An estimate of the accuracy of the identification of the entities and assignments of attribute values in the data set.

Attribute Accuracy Explanation

The identification of the test that yielded the Attribute Accuracy Value.

Positional Accuracy

Horizontal Accuracy

Horizontal Accuracy Report

An explanation of the accuracy of the horizontal coordinate measurements and a description of the tests used. For example, "Accuracy of these digital data (if not digitally revised), is based upon the use of source graphics which are compiled to meet National Map Accuracy Standards. NMAS horizontal accuracy requires that at least 90 percent of points tested are within 0.02 inches of the true position. The digital data are estimated to contain a horizontal positional error of less than or equal to 0.003 inches standard error in the two component directions relative to the source graphic. NMAS vertical accuracy requires that at least 90% of well defined points tested be within one half contour interval of the correct value...," "Horizontal positional accuracy for the digital data is tested by visual comparison of the source with hard copy plots," "unknown," or "to be determined."

OR

Quantitative Horiz. Position. Accuracy Assessment

Numeric value assigned to summarize the accuracy of the horizontal coordinate measurements and the identification of the test that yielded the value.

Horizontal Accuracy Value

An estimate of the accuracy of the horizontal coordinate measurements in the data set expressed in (ground) meters.

Horizontal Accuracy Explanation

The identification of the test that yielded the Horizontal Positional Accuracy Value.

Vertical Accuracy

Vertical Accuracy Report

An explanation of the accuracy of the vertical coordinate measurements and a description of the tests used.

OR

Quantitative Vert. Positional Accuracy Assessment

Vertical Accuracy Value

An estimate of the accuracy of the vertical coordinate measurements in the data set expressed in (ground) meters.

Vertical Accuracy Explanation

The identification of the test that yielded the Vertical Positional Accuracy Value.

Lineage

Information about the events, parameters, and source data which constructed the data set, and information about the responsible parties.

Source Information

List of sources and a short discussion of the information contributed by each.

Source Citation

Reference for a source data set.

Type of Source Media

The medium of the source data set. Examples include "paper," "stable-base material," "microfiche," "microfilm," "audiocassette," "chart," "filmstrip," "transparency," "videocassette," "videodisc," "videotape," "physical model," "computer program," "disc," "cartridge tape," "magnetic tape," "online," "CD-ROM," "electronic bulletin board," "electronic mail system," etc.

Format of Source Data

The file format of the source data set. Examples include “Tiff,” “Jpeg,” “.e00,” and other general and proprietary data formats.

Source Time Period of Content (YYYYMMDD)

Time period(s) for which the source data set corresponds to the ground, e.g., “calendar date: 1997,” etc.

Source Currentness Reference

The basis on which the source time period of content information of the source data set is determined. Examples include “source photography date,” “publication date,” etc.

Source Citation Abbreviation

Short-form alias for the source citation (several-letter abbreviation that can be used elsewhere in the metadata to refer to this source).

Source Contribution

Brief statement identifying the information contributed by the source to the data set.

Process Step

Process Description

An explanation of the activities (processes) performed on the data and related parameters or tolerances. Examples included projection, spatial analysis steps, etc.

Process Date (YYYYMMDD)

The date when the process was completed.

3. Spatial Reference

Coordinate System (auto filled if spatial reference has been applied to data)

Examples: “WGS84,” “Geographical,” etc.

Projection (auto filled if spatial reference has been applied to data)

E.g., “UTM Zone x,” “NM State Plane,” etc.

Datum (auto filled if spatial reference has been applied to data)

E.g., "NAD27," NAD83," etc.

OR

FIPS#

Standard reference number for projection/datum combination.

4. Entity and Attribute

The Overview Description would be used to describe data for which a complete description is available elsewhere; the Detailed Description would be used to describe data for which a complete description has not been compiled elsewhere. Refer to ArcCatalog Help for more information about entity and attributes descriptions for various data types (shapefiles, coverages, etc.).

Detailed Description

Description of the entities, attributes, attribute values, and related characteristics encoded in the data set. Users of a data set need to know the meaning of entity, attribute, and attribute value information associated with the spatial information. For example, a data set might include the entity "road". A "road" might have the attribute "road type," which can be assigned the attribute values of "heavy duty," "medium duty," "light duty," or "trail." The producer of the data set may have different definitions for "road," "road type," "heavy duty," "medium duty," "light duty," or "trail" than a user. The Entity and Attribute Information section provides the way for a producer to describe the meaning of this nonspatial entity, attribute, and attribute value information so a user can understand the information content of a data set and use the data appropriately.

Entity Type

Label (auto filled)

The name of the entity type, e.g., "wetland."

Definition

For example, "Wetlands are lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface…"

Definition Source

The authority (citation) of the definition given above.

Attribute Type

Label (auto filled)

The name of the attribute, e.g., "wetland classification."

Definition

The definition or background on the attribute label; for example, "the wetland classification system is hierarchical, with wetlands and deepwater habitats divided among five major systems…."

Definition Source

The authority (citation) of the definition given above.

AND/OR

Overview Description

In developing the metadata standards, reviewers noted that they often had complete data dictionaries for the entity, attribute, and attribute value information, and preferred to refer to the existing descriptions instead of duplicating them. The Overview Description provides the elements needed to give users a sense of the information content and a reference to the source(s) of the complete description. Note that the Overview Description does not relieve the producer of the responsibility to provide a complete description. The Overview Description elements refer to the source of the complete description, and do not allow a producer merely to summarize the information content.

Entity and Attribute Overview

Detailed summary of the information contained in a data set. For example, "The DLG-3 scheme assigns any number of attribute codes to spatial objects. Some DLG-3 attribute codes identify real world entities, while other codes represent attributes of those entities. Hydrography (major code: 050) contains information about streams, bodies of water, wetlands, coastal waters, and water used for transportation. Vegetative features, which are associated with wetlands or submerged areas, are included, as are manmade hydrologic features…."

Entity and Attribute Detail Citation

Reference to the complete description of the entity types, attributes, and attribute values for the data set. For example, "The USGS document Standards for Digital Line Graphs, Part 3: Attribute Codes, contains a detailed description of each attribute code and a reference to the associated map symbols on the map source materials…."

5. Distribution

Distribution Liability

Statement of the liability assumed by the distributor. GISLab uses the standard UC / DOE disclaimer: "Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or Implied, including the warranties of merchantability and fitness for a particular purpose, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights."

6. Metadata Reference

General

Metadata Date (auto filled)

Metadata Standard Name (auto filled)

Metadata Standard (auto filled)

Contact

The party responsible for the metadata information. The metadata contact should be the organization or person that can answer questions about the metadata or can receive reports about errors in the metadata.

Details

Person

Organization

General

Contact Voice Telephone

Address

Address Type

E.g., "mailing," "physical," "mailing and physical," etc.

Address

City

State or Province

Postal Code